



Remark on the finite-dimensional character of certain results of functional statistics

Jean-Marc Azais, Jean-Claude Fort

► To cite this version:

Jean-Marc Azais, Jean-Claude Fort. Remark on the finite-dimensional character of certain results of functional statistics. MAP5 2012-34. 4 pages. 2012. <hal-00759922>

HAL Id: hal-00759922

<https://hal.archives-ouvertes.fr/hal-00759922>

Submitted on 4 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Remark on the finite-dimensional character of certain results of functional statistics

Remarque sur le caractère fini dimensionnel de certains résultats de statistique fonctionnelle

Jean-Marc Azaïs ^{*} Jean-Claude Fort [†]

4 décembre 2012

Résumé

Cette note montre qu'une hypothèse concernant les probabilités de petites boules, fréquemment utilisée en statistique fonctionnelle, implique que la dimension de l'espace fonctionnel considéré est finie. Un exemple de processus L^2 , ne vérifiant pas cette hypothèse, vient compléter ce résultat.

Abstract

This note shows that some assumption on small balls probability, frequently used in the domain of functional statistics, implies that the considered functional space is of finite dimension. To complete this result an example of L^2 process is given that does not fulfill this assumption.

Mathematics Subject Classification : 60A10, 62G07

Keywords : functional statistics, finite dimension, small balls probability

1 The result

In several functional statistics papers (cf [1],[2],[3],[4],[5],[6]) the following hypothesis is used :

(H) *Let x be a point of the space \mathcal{X} where a functional variable X lives. The space \mathcal{X} is equipped with a semi distance and $B(x, h)$ is the ball with center x and radius $h > 0$. We set $\varphi_x(h) = \mathbb{P}(X \in B(x, h))$ and we assume :*

$$\inf_{h \in [0, 1]} \int_0^1 \varphi_x(ht) dt / \varphi_x(h) \geq \theta_x > 0.$$

where the parameter θ_x is locally bounded away from zero.

The aim of this note is to prove that (H) implies that \mathcal{X} is of finite dimension.

^{*}Université de Toulouse, IMT, ESP, F31062 Toulouse Cedex 9, France. Email : jean-marc.azais@math.univ-toulouse.fr

[†]Université Paris Descartes, SPC, MAP5, 45 rue des saints pères, F-75006 Paris; Email : jean-claude.fort@parisdescartes.fr

Proof : Without loss of generality we can assume that $\theta_x < 1/2$. Let $F_x(h) = \int_0^h \varphi_x(t)dt$ we have :

$$\frac{1}{h} \int_0^h \varphi_x(t)dt / \varphi_x(h) = \frac{F_x(h)}{hF'_x(h)} \geq \theta_x.$$

By integration we obtain

$$F_x(h)/F_x(1) \geq h^{1/\theta_x}. \quad (1)$$

Since φ is non decreasing we have

$$F_x(h) \leq h\varphi_x(h)$$

and thus

$$\varphi_x(h) \geq h^{1/\theta_x-1}F_x(1).$$

Let $x \in \mathcal{X}$ such that the parameter θ_y has positive lower bound for y in the ball $B(x, h_0)$. This implies that $\varphi_x(h_0) > 0$. By a scaling we can assume for simplicity that $h_0 = 1$. For all y in $B(x, 1/4)$, and for all $h \in [1/2, 1]$,

$$\varphi_y(h) \geq \varphi_x(1/4) \geq \left(\frac{1}{4}\right)^{1/\theta-1}F_x(1),$$

where θ is the uniform lower bound for θ_y for y in $B(x, 1)$. By integration

$$F_y(1) \geq 1/2\left(\frac{1}{4}\right)^{1/\theta-1}F_x(1)$$

and

$$\varphi_y(h) \geq h^{1/\theta-1}F_y(1) \geq 1/2\left(\frac{1}{4}\right)^{1/\theta-1}h^{1/\theta-1}F_x(1)$$

This implies that there exist at most $\mathcal{O}(h^{1-\theta})$ disjoint ball of radius h in $B(x, 1/4)$. Then the same set of balls but with radius $2r$ is a covering, which implies in turn that the box (or entropy) dimension of $B(x, 1/4)$ is finite. Since the Hausdorff dimension is smaller than the box dimension it is also finite.

Remark 1. Suppose, in addition, that the probability distribution function of X satisfying H admits a density with respect to some natural positive measure μ and suppose that this density is locally upper- and lower -bounded by M and m respectively :

$$m\mu(B(x, h)) \leq \varphi_x(h) \leq M\mu(B(x, h)).$$

We have then

$$\exists d > 0, C_2 > 0, h_0 > 0, \quad 0 \leq h \leq h_0 \implies 1/C_2 h^d \leq \mu(B(x, h)) \leq C_2 h^d.$$

which means that μ shares the same property as the distribution of X . As for example this property does not hold for the Wiener measure.

We may note that the classical random processes : Brownian motion and more general Gaussian processes for which the probability of small balls is known do not satisfy H , see [7] for more details.

2 Random series of functions

Here we consider $X(t)$ a stochastic process constructed as a random series of functions. More precisely

$$X(t) = \sum_{n=1}^{\infty} \alpha_n Z_n \varphi_n(t) \quad (2)$$

where

- $(Z_n, n \geq 1)$ is a sequence of independent real random variables, with mean 0 and variance 1, with absolutely continuous densities $(f_n, n \geq 1)$ w.r.t to the Lebesgue measure. This is the case for most of the usual continuous distributions of probability : exponential, normal, polynomial, gamma, beta etc...
 - $(1; \varphi_n, n \geq 1)$ is an orthonormal basis of $L^2([0, 1])$.
 - $(\alpha_n, n \geq 1)$ is a sequence of positive real numbers $\sum_{n=1}^{+\infty} \alpha_n^2 < \infty$.
- The sum (2) converges in $L^2([0, 1])$.

In particular the form (2) covers all the Gaussian processes through the Karhunen-Loève decomposition.

Then we have :

Proposition 2.1. $\lim_{h \rightarrow 0} h^d \mathbb{P}(\|X\|_2 \leq h) = 0$ for any $d \geq 0$. So that the process $X(t)$ does not fulfill the assumption (H) for the L^2 norm at the point zero.

The proof is based on the properties of the convolution :

Let f and g the absolutely continuous densities of probability of two independent random variables U and V . Then U^2 (resp. V^2) has the density

$$p_{U^2}(u) = \frac{\tilde{f}(\sqrt{u})}{\sqrt{u}} \quad \text{resp.} \quad p_{V^2}(v) = \frac{\tilde{g}(\sqrt{v})}{\sqrt{v}},$$

where \tilde{f} and \tilde{g} are the symmetrized of f and g , $\tilde{f} = 1/2(f(x) + f(-x))$. It follows that $U^2 + V^2$ has for density

$$C(x) = \int_0^x \frac{\tilde{f}(\sqrt{u})}{\sqrt{u}} \frac{\tilde{g}(\sqrt{x-u})}{\sqrt{x-u}} du = \int_0^1 \frac{1}{\sqrt{v(1-v)}} \tilde{g}(\sqrt{(1-v)x}) \tilde{f}(\sqrt{vx}) dv,$$

for $x \geq 0$ and 0 elsewhere.

It is easy to see that for any $0 < A < B$, the function C is Lipschitz on $[A, B]$. At $x = 0$ it takes the value 0 with right limit $\beta(\frac{1}{2}, \frac{1}{2})f(0)g(0)$. Now if we make the convolution product of two such functions C_1 and C_2 , we obtain a function vanishing for $x \leq 0$, continuous at 0 and Lipschitz on any compact interval of \mathbb{R} , thus absolutely continuous. Using a classical result, making the convolution product of k such absolutely continuous functions yields a C^{k-1} function whose $(k-1)^{\text{th}}$ derivative is absolutely continuous.

Then, applying iteratively this result we conclude that the density of the variable $\|X\|_2^2 = \sum_{n=1}^{\infty} \alpha_n^2 Z_n^2$ is infinitely derivable with all its derivatives null at 0. The claimed property follows.

Remark 2. If the process X lives in $L^p[0, 1]$, $p \in (2, \infty]$, we have $\|X\|_2^2 \leq \|X\|_p^2$, thus :

$$\mathbb{P}(\|X\|_p \leq h) \leq \mathbb{P}(\|X\|_2 \leq h).$$

References

- [1] Burba, Florent ; Ferraty, Frédéric ; Vieu, Philippe. k -nearest neighbour method in functional nonparametric regression. J. Nonparametr. Stat. 21 (2009), no. 4, 453–469.
- [2] Dabo-Niang, Sophie ; Ferraty, Frédéric ; Vieu, Philippe. Estimation du mode dans un espace vectoriel semi-normé. (French) [Mode estimation in a semi-normed vector space] C. R. Math. Acad. Sci. Paris 339 (2004), no. 9, 659–662.
- [3] Ferraty, Frédéric ; Kudraszow, Nadia ; Vieu, Philippe Nonparametric estimation of a surrogate density function in infinite-dimensional spaces. J. Nonparametr. Stat. 24 (2012), no. 2, 447–464.
- [4] Ferraty, F. ; Van Keilegom, I. ; Vieu, P. Regression when both response and predictor are functions. J. Multivariate Anal. 109 (2012), , 10–28.
- [5] Ferraty, Frédéric ; Laksaci, Ali ; Tadj, Amel ; Vieu, Philippe. Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plann. Inference 140 (2010), no. 2, 335–352.
- [6] Ferraty, Frédéric ; Goia, Aldo ; Vieu, Philippe. Régression non-paramétrique pour des variables aléatoires fonctionnelles mélangeantes. C. R. Math. Acad. Sci. Paris 334 (2002), no. 3, 217–220.
- [7] Li, Wenbo V. ; Shao, Qi-Man. Gaussian Processes : Inequalities, Small Ball Probabilities and Applications. Stochastic processes : Theory and methods, Handbook of Statistics, Vol. 19, Edited by C.R. Rao and D. Shanbhag, Elsevier, New York (2001), 533–598.